

# A NOTE ON THE FINITE-DIMENSIONAL DISTRIBUTIONS OF DISPERSING BILLIARD PROCESSES

JUHO LEPPÄNEN AND MIKKO STENLUND

**ABSTRACT.** In this short note we consider the finite-dimensional distributions of sets of states generated by dispersing billiards with a random initial condition. We establish a functional correlation bound on the distance between the finite-dimensional distributions and corresponding product distributions. We demonstrate the usefulness of the bound by showing that it implies several limit theorems. The purpose of this note is to provide a tool facilitating the study of more general functionals of the billiard process.

**Acknowledgements.** This work was supported by the Jane and Aatos Erkkö Foundation, and by Emil Aaltosen Säätiö.

## 1. INTRODUCTION

In this note we revisit the two-dimensional dispersing Sinai billiard with finite horizon. To specify the model, we consider the torus  $\mathbb{T}^2$  with a finite collection of scatterers, i.e., closed convex sets  $S_1, \dots, S_M$ , having  $C^3$  boundaries  $\partial S_m$  with strictly positive curvatures. A particle moves linearly in the domain  $\mathbb{T}^2 \setminus \cup_{m=1}^M S_m$ , with unit speed, up to elastic collisions with the boundaries  $\partial S_m$  of the scatterers. The scatterers are disjoint and positioned so that the free path length of the particle is bounded. Note that the precise dynamics of the system is fully determined by the geometry of the domain  $\mathbb{T}^2 \setminus \cup_{m=1}^M S_m$ .

A standard discrete-time representation of the dynamics is obtained by keeping track of the collisions only, which leads to the so-called collision map  $T : X \rightarrow X$  as follows: Topologically,  $X$  is the disjoint union of  $m$  cylinders  $X_m$ , homeomorphic to  $\partial S_m \times [-\frac{\pi}{2}, \frac{\pi}{2}]$ . A general point  $x \in X$  consists of a pair  $x = (r, \varphi)$ , where  $r$  represents the position of the particle on the boundary  $\cup_{m=1}^M \partial S_m$  during a collision, and  $\varphi$  represents its direction immediately after the collision relative to the normal of the boundary. Then  $Tx = (r', \varphi')$  is defined as the corresponding pair after the next collision. Since the continuous-time system is Hamiltonian, preserving phase-space volume, the collision map preserves a corresponding Borel probability measure, namely  $d\mu(r, \varphi) = \text{const} \cdot \cos \varphi \, dr \, d\varphi$ , on  $X$ . Reversing the velocity of the particle, one moreover verifies that  $T$  is invertible.

Given an initial state  $x \in X$ , the billiard dynamics generates the sequence of states  $(T^i x)_{i \in \mathbb{N}} \in X^{\mathbb{N}}$ . If  $x$  is chosen at random, according to the invariant measure  $\mu$ , the sequence is a stationary random process, which we call the billiard process. We can equally well define the two-sided billiard process  $(T^i x)_{i \in \mathbb{Z}}$ , and everything below extends readily to that setup, but let us proceed with the one-sided case. Of course, knowledge of the value of  $T^i x$  for some  $i$  fully determines the value of  $T^j x$  for all  $j$ . Yet, the same is in general not true of, say,  $f(T^i x)$  and  $g(T^j x)$  where  $f, g : X \rightarrow \mathbb{R}$  are “observables”. For instance, if  $f$  and  $g$  are Hölder continuous, then an

---

*Key words and phrases.* Dispersing billiards, finite-dimensional distributions, weak convergence, decorrelation.

exponential covariance bound

$$\left| \int f(T^i x) g(T^j x) d\mu(x) - \int f(T^i x) d\mu(x) \int g(T^j x) d\mu(x) \right| \leq C\theta^{j-i} \quad (1)$$

holds. Above,  $C > 0$  and  $\theta \in (0, 1)$  depend on the Hölder classes of  $f, g$  and on system constants<sup>1</sup>. Obviously, more general bounds exist, but (1) is sufficient for the ongoing illustrative discussion. The covariance bound (1), and other similar results, are consequences of the chaotic nature of the billiard dynamics. Colloquially, we may regard the observations  $f(T^i x)$  as weakly dependent random variables, which in one form or another is at the heart of proving probabilistic limit results for functionals  $F \circ (T^i x)_{i \in \mathbb{N}}$  of the billiard process, say concerning the asymptotic behavior of the Birkhoff sums

$$S_N(x) = \sum_{i=0}^{N-1} f(T^i x)$$

in the limit  $N \rightarrow \infty$ .

To proceed, we recall that a finite-dimensional distribution of the billiard process is the joint distribution  $\mathbf{P}_I$  of a subsequence  $(T^i(x))_{i \in I}$  corresponding to a finite index set  $I = (i_1, i_2, \dots, i_p) \subset \mathbb{N}$ . From here on, we will without loss of generality always assume that the indices in such an index set are in increasing order,  $i_1 < \dots < i_p$ . The probability measure  $\mathbf{P}_I$  on  $X^p$  is characterized by the identity

$$\int h d\mathbf{P}_I = \int h(T^{i_1} x, T^{i_2} x, \dots, T^{i_p} x) d\mu(x)$$

for bounded measurable functions  $h : X^p \rightarrow \mathbb{R}$ . Of course, stationarity of the billiard process means that  $\mathbf{P}_I = \mathbf{P}_{I'}$  for all translates  $I' = (i_1 + k, \dots, i_p + k)$  of  $I$ , which is clearly true by the invariance of  $\mu$ .

For example, in terms of finite-dimensional distributions, (1) reads

$$\left| \int f \otimes g d\mathbf{P}_{(i,j)} - \int f \otimes g d(\mathbf{P}_i \otimes \mathbf{P}_j) \right| \leq C\theta^{j-i}$$

where  $(f \otimes g)(x, y) = f(x)g(y)$ . In this weak sense, we may informally write

$$\text{“ } \mathbf{P}_{(i,j)} \approx \mathbf{P}_i \otimes \mathbf{P}_j \text{ ”}$$

when  $j - i$  is large.

**Convention.** In the rest of the note we will consider unions

$$I = \bigcup_{1 \leq k \leq K} I_k$$

of increasing nonempty index sets  $I_k = (i_{p_{k-1}+1}, \dots, i_{p_k}) \subset \mathbb{N}$ , where  $p_0 = 0$ . We will always assume they are disjoint and ordered, in the sense that the gap between  $I_k$  and  $I_{k+1}$  satisfies

$$\ell_k = i_{p_k+1} - i_{p_k} > 0$$

for all  $k = 1, \dots, K - 1$ . We shall henceforth write

$$I_1 \leq \dots \leq I_K$$

to express these conventions succinctly.

Being still informal, higher order correlation bounds indicate that, when each  $\ell_k$  is large,

$$\text{“ } \mathbf{P}_I \approx \mathbf{P}_{I_1} \otimes \dots \otimes \mathbf{P}_{I_K} \text{ ”} \quad (2)$$

---

<sup>1</sup>In this paper system constants are quantities which only depend on the geometry of the domain  $\mathbb{T}^2 \setminus \bigcup_{m=1}^M S_m$ .

in some weak sense. The purpose of this brief note is to make this interpretation precise. As an aside, it provides a unified perspective on several limit theorems that we treat as examples: We will obtain an estimate on the difference between  $\mathbf{P}_I$  and  $\mathbf{P}_{I_1} \otimes \cdots \otimes \mathbf{P}_{I_K}$  in an appropriately general sense of practical use, which is then shown to imply all the limit theorems. Let us immediately be clear that the latter limit theorems, per se, have been proved elsewhere, in the references cited (although we do obtain some minor improvements). Thus a side goal here is to shed additional light on why those theorems are true, in a mathematically rigorous way. The main result is the aforementioned estimate, which we call the “functional correlation bound”. We expect it to be of much broader use, as it is directly applicable to studying other kinds of functionals of the billiard process than the examples included here. In short, we view the functional correlation bound as a tool which helps put the vague statement in (2) onto a solid footing, in reasonable generality, so that it can be used effectively in technical proofs in the theory of dispersing billiards.

**Structure of this note.** In Section 2 we state two theorems on functional correlation decay. In Section 3 we give examples of using them for deducing limit theorems, and in Section 4 we prove them.

## 2. RESULTS

Before stating the results, we recall a few standard facts from the theory of dispersing billiards. The reader is referred to the book [3] for more details.

In the disjoint union  $X = \cup_{m=1}^M X_m$ , the cylinders  $X_m = \partial S_m \times [-\frac{\pi}{2}, \frac{\pi}{2}]$  are further divided into horizontal strips  $X_{m,k} = \partial S_m \times \{b_k < \varphi < b_{k+1}\}$ ,  $k \in \mathbb{Z}$ , called homogeneity strips. (Here the numbers  $b_k$  are such that  $b_k - b_{k+1} = O(k^{-3})$ , which facilitates controlling distortions of the map  $T$  within each strip.) We consider the totality of the homogeneity strips the connected components of  $X$ . For a pair of points  $x, y \in X$ , we say that their trajectories separate when  $T^n x$  and  $T^n y$  are in different components for the first time  $n \geq 0$ ; this  $n$  is called the future separation time, which we denote by  $s_+(x, y)$ . We define  $s_+(x, y) = \infty$  if the trajectories never separate. The past separation time  $s_-(x, y)$  is the analogous notion for the inverse map  $T^{-1}$ .

A local stable manifold  $W^s(x)$  of a point  $x \in X$  is a maximal  $C^2$  curve such that  $T^n W^s(x)$  is completely contained in a component of  $X$ , for all  $n \geq 0$ . That is, given  $n \geq 0$ , there exist  $m, k$  such that  $T^n W^s(x) \subset X_{m,k}$ . It can be shown that almost every point has a nontrivial local stable manifold, and that the length of  $T^n W^s(x)$  decreases exponentially as  $n \rightarrow \infty$ . Given two points  $x, y \in X$ , we either have  $W^s(x) = W^s(y)$  (meaning  $y \in W^s(x)$ ) or  $W^s(x) \cap W^s(y) = \emptyset$ . Note that, in the first case,  $s_-(T^n x, T^n y) = n + s_-(x, y)$  for all  $n \geq 0$ . The family of all local stable manifolds is uncountable, and forms a measurable partition of  $X$ . Local unstable manifolds  $W^u(x)$  have identical properties in terms of the inverse map  $T^{-1}$ . In particular, they form a measurable partition of  $X$ . Moreover, if  $y \in W^u(x)$ , then  $s_+(T^{-n} x, T^{-n} y) = n + s_+(x, y)$  for all  $n \geq 0$ .

We also recall the notion of dynamical Hölder continuity. The following definition is from [15]. It is a small variation of the one in [2], but in the current form it enjoys the property of being dynamically closed, which is used in the proofs; see Lemma 4.5.

**Definition 2.1.** A function  $g : X \rightarrow \mathbb{R}$  is *dynamically Hölder continuous on local unstable manifolds* with rate  $\vartheta \in (0, 1)$  and constant  $c \geq 0$  if

$$|g(x) - g(y)| \leq c \vartheta^{s_+(x, y)}$$

holds whenever  $x$  and  $y$  belong to the same local unstable manifold. In this case we write  $g \in \mathcal{H}_+(c, \vartheta)$ . Likewise,  $g$  is dynamically Hölder continuous on local stable manifolds if

$$|g(x) - g(y)| \leq c \vartheta^{s-(x,y)}$$

holds whenever  $x$  and  $y$  belong to the same local stable manifold. In this case we write  $g \in \mathcal{H}_-(c, \vartheta)$ .

For instance, if  $g : X \rightarrow \mathbb{R}$  is Hölder continuous with exponent  $\alpha \in (0, 1)$  and constant  $|g|_\alpha$ , then  $g \in \mathcal{H}_-(c, \vartheta) \cap \mathcal{H}_+(c, \vartheta)$ , where  $c = \text{const} \cdot |g|_\alpha$ , and  $\vartheta = \vartheta(\alpha)$  is determined by  $\alpha$  and system constants.

Finally, we introduce the class of admissible test functions  $F$ :

**Definition 2.2.** Given increasingly ordered index sets  $I_1 \leq \dots \leq I_K$ ,  $K \geq 2$ , we say that a bounded function  $F : X^{p_K} \rightarrow \mathbb{R}$  is  $(I_1, \dots, I_K)$ -admissible, if it is separately dynamically Hölder continuous in the sense that

$$x_r \mapsto F(x_1, \dots, x_{p_K}) \in \begin{cases} \mathcal{H}_+(c_r, \vartheta_+), & 1 \leq r \leq p_1, \\ \mathcal{H}_+(c_{r+}, \vartheta_+) \cap \mathcal{H}_-(c_{r-}, \vartheta_-), & p_1 + 1 \leq r \leq p_{K-1}, \\ \mathcal{H}_-(c_r, \vartheta_-), & p_{K-1} + 1 \leq r \leq p_K. \end{cases}$$

Here is our first functional correlation bound, concerning the case  $K = 2$ :

**Theorem 2.3.** There exist system constants  $M_0, M_1 > 0$  and  $\theta_0, \theta_1 \in (0, 1)$  such that the following holds. Let  $I_1 \leq I_2$ , and let  $F$  be  $(I_1, I_2)$ -admissible:

$$x_r \mapsto F(x_1, \dots, x_{p_2}) \in \begin{cases} \mathcal{H}_+(c_r, \vartheta_+), & 1 \leq r \leq p_1, \\ \mathcal{H}_-(c_r, \vartheta_-), & p_1 + 1 \leq r \leq p_2. \end{cases}$$

Then

$$\begin{aligned} & \left| \int F d\mathbf{P}_{I_1 \cup I_2} - \int F d(\mathbf{P}_{I_1} \otimes \mathbf{P}_{I_2}) \right| \\ & \leq \left( \sum_{r=1}^{p_1} c_r \vartheta_+^{i_{p_1} - i_r} \right) \vartheta_+^{\frac{1}{4}\ell_1} + M_0 \left( \sum_{r=p_1+1}^{p_2} c_r \vartheta_-^{i_r - i_{p_1+1}} + \|F\|_\infty \right) \max(\theta_0, \vartheta_-)^{\frac{1}{4}\ell_1 - \frac{1}{3}} \\ & \quad + 2M_1 \|F\|_\infty \theta_1^{\frac{1}{4}\ell_1 - \frac{1}{3}}. \end{aligned}$$

Here  $\ell_1 = i_{p_1+1} - i_{p_1}$  is the gap between  $I_1$  and  $I_2$ .

The second functional correlation bound extends the first one to  $K \geq 2$ . While it is entirely possible to formulate the result for general admissible test functions  $F$ , the resulting bound has a cumbersome expression. For aesthetic reasons alone, we restrict to functions admissible with the same parameters, leaving generalizations to the reader. By “the same parameters” we mean that  $c_r \equiv c_{r\pm} \equiv c$  and  $\vartheta_- = \vartheta_+$  in Definition 2.2.

**Theorem 2.4.** Let  $I_1 \leq \dots \leq I_K$ ,  $K \geq 2$ , and let  $F$  be  $(I_1, \dots, I_K)$ -admissible, with the same parameters  $c \geq 0$  and  $\vartheta \in (0, 1)$ . Then

$$\left| \int F d\mathbf{P}_I - \int F d(\mathbf{P}_{I_1} \otimes \dots \otimes \mathbf{P}_{I_K}) \right| \leq M \left( \frac{c}{1 - \vartheta} + \|F\|_\infty \right) \sum_{k=1}^{K-1} \theta^{\ell_k}.$$

Here  $\ell_k = i_{p_k+1} - i_{p_k}$  is the gap between  $I_k$  and  $I_{k+1}$ ,

$$\theta = \max(\vartheta, \theta_0, \theta_1)^{\frac{1}{4}}$$

and

$$M = \max\left(1 + M_0\theta_0^{-\frac{1}{3}}, M_0\theta_0^{-\frac{1}{3}} + 2M_1\theta_1^{-\frac{1}{3}}\right).$$

The system constants  $M_0, M_1$  and  $\theta_0, \theta_1$  are the same as in Theorem 2.3.

A result in the spirit of Theorem 2.4 was recently proved by Leppänen [7], for a class of one-dimensional, non-uniformly expanding dynamical systems.

In fact, the inductive proof of Theorem 2.4 shows that the special case  $K = 2$  and the general case  $K \geq 2$  are equivalent. This hinges on the dynamical closedness of the function classes  $\mathcal{H}_-$  and  $\mathcal{H}_+$ ; see Lemma 4.5.

At first, the theorems may seem like inconsequential extensions of correlation bounds such as the one displayed in (1). But they do allow for estimating integrals of functionals of the billiards process,  $\int F \circ (T^i x)_{i \in \mathbb{N}} d\mu$ , beyond the scope of simple correlation bounds. Just to give a simple example, consider a situation of the following kind:

**Example 2.5.** Let  $A : \mathbb{R}^K \rightarrow \mathbb{R}$  be Lipschitz continuous with constant  $L$  in each variable, and let the index sets  $I_1 \leq \dots \leq I_K$  be as above. Define the sums

$$S^{(k)} = \sum_{r=p_{k-1}+1}^{p_k} f_r \circ T^{i_r}, \quad 1 \leq k \leq K,$$

where the functions  $f_r : X \rightarrow \mathbb{R}$  are bounded, and dynamically Hölder continuous with the same parameters  $c \geq 0$  and  $\vartheta \in (0, 1)$ . Let us consider the integral

$$\mathcal{J} = \int A(S^{(1)}(x), \dots, S^{(K)}(x)) d\mu(x).$$

We would like to argue that, when each  $\ell_k = i_{p_k+1} - i_{p_k}$  is large, the sums in the argument of  $A$  are weakly dependent, so  $\mathcal{J}$  must be close to

$$\mathcal{J}^\otimes = \int A(S^{(1)}(x_1), \dots, S^{(K)}(x_K)) d\mu^{\otimes K}(x_1, \dots, x_K),$$

where the sums are literally independent due to the product measure. Theorem 2.4 helps make such an argument rigorous: Let  $F : X^{p_K} \rightarrow \mathbb{R}$  be the function

$$F(x_1, \dots, x_{p_K}) = A\left(\sum_{r=1}^{p_1} f_r(x_r), \dots, \sum_{r=p_{K-1}+1}^{p_K} f_r(x_r)\right).$$

Then  $F$  is  $(I_1, \dots, I_K)$ -admissible:  $x_r \mapsto F(x_1, \dots, x_{p_K}) \in \mathcal{H}_+(Lc, \vartheta) \cap \mathcal{H}_-(Lc, \vartheta)$  for all  $r$ , and

$$\|F\|_\infty \leq C = \begin{cases} \|A\|_\infty & \text{if } A \text{ is bounded,} \\ L \sum_{r=1}^{p_K} \|f_r\|_\infty + A(0, \dots, 0) & \text{if } A \text{ is unbounded.} \end{cases}$$

In both cases, we immediately arrive at the quantitative estimate

$$|\mathcal{J} - \mathcal{J}^\otimes| \leq M \left( \frac{Lc}{1 - \vartheta} + C \right) \sum_{k=1}^{K-1} \theta^{\ell_k}.$$

For instance, Example 2.5 applies to “interlaced” covariances of the form

$$\begin{aligned} & \int A_1(S^{(1)} + S^{(3)} + \dots + S^{(K-1)}) A_2(S^{(2)} + S^{(4)} + \dots + S^{(K)}) d\mu \\ & - \int A_1(S^{(1)} + S^{(3)} + \dots + S^{(K-1)}) d\mu \int A_2(S^{(2)} + S^{(4)} + \dots + S^{(K)}) d\mu, \end{aligned}$$

where the argument of the Lipschitz function  $A_1$  (respectively,  $A_2$ ) involves  $f_r \circ T^{i_r}$  with  $i_r \in I_k$  and  $k$  odd (respectively,  $k$  even). Here  $K$  is even for convenience. This is so, because both terms in the difference can be compared with the integral with respect to the product measure  $d\mu^{\otimes K}$ .

In the special case of singleton index sets,  $I_k = \{i_k\}$ , Example 2.5 yields a bound on

$$\int A(f_1(T^{i_1}x), \dots, f_K(T^{i_K}x)) d\mu(x) - \int A(f_1(T^{i_1}x_1), \dots, f_K(T^{i_K}x_K)) d\mu^{\otimes K}(x_1, \dots, x_K).$$

Such bounds are relevant, e.g., for multiple recurrence problems.

### 3. MORE EXAMPLES

In this section we give applications of Theorem 2.4 to limit results. We reiterate that the sole purpose of this section is to illustrate the usefulness of the theorem: it allows to check the conditions of various limit theorems with great ease. The verified conditions actually amount to *very* simple special cases of Theorem 2.4. We thus believe the result to be a tool of much broader use in analyzing dispersing billiard dynamics.

Below, the various constants in the results concerning billiards will be the same as in Theorem 2.4.

#### 3.1. Multiple correlation bounds.

**Theorem 3.1.** *Let  $f_0, \dots, f_r \in \mathcal{H}_+(c, \vartheta)$  and  $g_0, \dots, g_k \in \mathcal{H}_-(c, \vartheta)$ , and define*

$$\tilde{f} = f_0 \cdot f_1 \circ T^{-1} \cdots f_r \circ T^{-r} \quad \text{and} \quad \tilde{g} = g_0 \cdot g_1 \circ T^1 \cdots g_k \circ T^k.$$

*Suppose  $\|f_u\|_\infty = \max_{0 \leq i \leq r} \|f_i\|_\infty$  and  $\|g_v\|_\infty = \max_{0 \leq i \leq k} \|g_i\|_\infty$ . Then*

$$\left| \int \tilde{f} \cdot \tilde{g} \circ T^n d\mu - \int \tilde{f} d\mu \int \tilde{g} d\mu \right| \leq C\theta^n$$

*for all  $n \geq 0$ , where*

$$C = M\|f_u\|_\infty^r \|g_v\|_\infty^k \max\{\|f_u\|_\infty, \|g_v\|_\infty\} \left( \frac{c}{1-\vartheta} + \min\{\|f_u\|_\infty, \|g_v\|_\infty\} \right).$$

It was shown in [2] that such a multiple correlation bound suffices for the central limit theorem to hold: If  $f \in \mathcal{H}_-(c, \vartheta) \cap \mathcal{H}_+(c, \vartheta)$  is bounded and  $\int f d\mu = 0$ , then

$$\frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} f \circ T^i$$

converges weakly, as  $N \rightarrow \infty$ , to the normal distribution  $\mathcal{N}(0, \sigma^2)$  with zero mean and variance

$$\sigma^2 = \int f^2 d\mu + 2 \sum_{i=1}^{\infty} \int f f \circ T^i d\mu.$$

See also [10] for a closely related result. To be technically accurate, [2] dealt with a smaller class of observables, as did [10]. In [15] it was shown that, for the present classes  $\mathcal{H}_\pm$ , the multiple correlation bound is equivalent to the pair correlation bound corresponding to the special case  $r = k = 0$ ; and consequently that the pair correlation bound alone is enough for the central limit theorem.

*Proof of Theorem 3.1.* Define  $I_1 = (0, \dots, r)$  and  $I_2 = (n + r, \dots, n + r + k)$ . Then

$$\int \tilde{f} \cdot \tilde{g} \circ T^n d\mu = \int F d\mathbf{P}_{I_1 \cup I_2} \quad \text{and} \quad \int \tilde{f} d\mu \int \tilde{g} d\mu = \int F d(\mathbf{P}_{I_1} \otimes \mathbf{P}_{I_2})$$

where  $F : X^{r+k+2} \rightarrow \mathbb{R}$  is the function

$$F(x_1, \dots, x_{r+1}, x_{r+2}, \dots, x_{r+k+2}) = f_r(x_1) \cdots f_0(x_{r+1}) g_0(x_{r+2}) \cdots g_k(x_{r+k+2}).$$

Set  $c_F = c \|f_u\|_\infty^r \|g_v\|_\infty^k \max\{\|f_u\|_\infty, \|g_v\|_\infty\}$ . Since  $f_i \in \mathcal{H}_+(c, \vartheta)$  and  $g_i \in \mathcal{H}_-(c, \vartheta)$ , we have

$$x_j \mapsto F(x_1, \dots, x_{r+k+2}) \in \begin{cases} \mathcal{H}_+(c_F, \vartheta), & 1 \leq j \leq r+1, \\ \mathcal{H}_-(c_F, \vartheta), & r+2 \leq j \leq r+k+2. \end{cases}$$

Hence,  $F$  is  $(I_1, I_2)$ -admissible with the same parameters  $c_F$  and  $\vartheta$ . By Theorem 2.4,

$$\left| \int \tilde{f} \cdot \tilde{g} \circ T^n d\mu - \int \tilde{f} d\mu \int \tilde{g} d\mu \right| \leq M \left( \frac{c_F}{1 - \vartheta} + \|F\|_\infty \right) \theta^n,$$

which implies the desired bound.  $\square$

**3.2. Multivariate normal approximation by Stein's method.** In this section and the next, we show that Theorem 2.4 implies not only normal convergence but also estimates on the speed of convergence. In particular, we treat the case of multivariate normal distributions arising from vector-valued observables.

Let  $T : X \rightarrow X$  be a general transformation preserving a probability measure  $\mu$ . We introduce the following notations: Given an observable  $f : X \rightarrow \mathbb{R}^d$ , we write

$$f^k = f \circ T^k$$

for all  $k \geq 0$ , denoting the coordinate functions of  $f^k$  by  $f_\alpha^k$ ,  $\alpha \in \{1, \dots, d\}$ . We set

$$W_N = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f^k$$

for all  $N \geq 0$ . For  $0 \leq K < N$ , we introduce the time window

$$[n]_{N,K} = \{k \in \mathbb{N}_0 \cap [0, N-1] : |k - n| \leq K\}$$

around  $n \geq 0$ , and define

$$W_{N,K}^n = W_N - \frac{1}{\sqrt{N}} \sum_{k \in [n]_{N,K}} f^k$$

for  $0 \leq n < N-1$ . Thus,  $W_{N,K}^n$  is a modification of  $W_N$  where the times  $k \in [n]_{N,K}$  are omitted in the sum. Finally,  $\Phi_\Sigma(h)$  stands for the expectation of a function  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  with respect to the centered multivariate normal distribution with covariance matrix  $\Sigma$ . We write  $\|f\|_\infty = \max_{1 \leq \alpha \leq d} \|f_\alpha\|_\infty$ ,  $\|D^2 h\|_\infty = \max_{1 \leq \alpha, \beta \leq d} \|\partial_\alpha \partial_\beta h\|_\infty$ , etc., for the norms of tensor fields.

The following theorem was proved in [5], where an adaptation of Stein's method [14] to the study of dynamical systems was developed:

**Theorem 3.2.** *Let  $f : X \rightarrow \mathbb{R}^d$  be a bounded measurable function with  $\mu(f) = 0$ . Suppose  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  is three times differentiable with  $\|D^k h\|_\infty < \infty$  for  $1 \leq k \leq 3$ . Fix integer  $N > 2$ . Suppose there exists  $\theta \in (0, 1)$  such that the following conditions are satisfied:*



(A1) There exist constants  $C_2 > 0$  and  $C_4 > 0$  such that

$$\begin{aligned} |\mu(f_\alpha f_\beta^k)| &\leq C_2 \theta^k \\ |\mu(f_\alpha f_\beta^l f_\gamma^m f_\delta^n)| &\leq C_4 \min\{\theta^l, \theta^{n-m}\} \\ |\mu(f_\alpha f_\beta^l f_\gamma^m f_\delta^n) - \mu(f_\alpha f_\beta^l) \mu(f_\gamma^m f_\delta^n)| &\leq C_4 \theta^{m-l} \end{aligned}$$

hold whenever  $k \geq 0$ ;  $0 \leq l \leq m \leq n < N$ ; and  $\alpha, \beta, \gamma, \delta \in \{1, \dots, d\}$ .

(A2) There exists a constant  $C_0$  such that

$$|\mu(f^n \cdot \nabla h(v + W_{N,K}^n t))| \leq C_0 \theta^K$$

holds for all  $0 \leq n < N$ ,  $0 \leq t \leq 1$ ,  $v \in \mathbb{R}^d$ , and  $0 \leq K < N$ .

(A3)  $f$  is not a coboundary in any direction.<sup>2</sup>

Then

$$\Sigma = \mu(f \otimes f) + \sum_{n=1}^{\infty} (\mu(f^n \otimes f) + \mu(f \otimes f^n)) \quad (3)$$

is a well-defined, symmetric, positive-definite,  $d \times d$  matrix; and

$$|\mu(h(W_N)) - \Phi_\Sigma(h)| \leq \left( C_* \left( \frac{2}{|\log \theta|} + \frac{\theta}{\sqrt{3}(1-\theta)} \right) + C_0 \right) \cdot \frac{\log N}{\sqrt{N}} \quad (4)$$

where

$$C_* = 12d^3 \max\{C_2, \sqrt{C_4}\} (\|D^2 h\|_\infty + \|f\|_\infty \|D^3 h\|_\infty) \sum_{i=0}^{\infty} (i+1) \theta^i.$$

Returning to billiards, we now prove the following:

**Theorem 3.3.** Assume  $f : X \rightarrow \mathbb{R}^d$  is bounded,  $\int f d\mu = 0$ , and there exist constants  $c \geq 0$  and  $\vartheta \in (0, 1)$  such that  $f_\alpha \in \mathcal{H}_-(c, \vartheta) \cap \mathcal{H}_+(c, \vartheta)$  for all  $\alpha \in \{1, \dots, d\}$ . Then, for all  $N$ , condition (A1) is satisfied with

$$C_2 = M \|f\|_\infty \left( \frac{c}{1-\vartheta} + \|f\|_\infty \right) \quad \text{and} \quad C_4 = M \|f\|_\infty^3 \left( \frac{c}{1-\vartheta} + \|f\|_\infty \right),$$

and condition (A2) is satisfied with

$$C_0 = M \left( d^2 c \frac{\|\nabla h\|_\infty + \|f\|_\infty \|D^2 h\|_\infty}{1-\vartheta} + d \|f\|_\infty \|\nabla h\|_\infty \right).$$

A similar result (with different constants) was recently proved in [5], but there a direct scheme for checking (A2) was implemented. Here we illustrate that (A1) and (A2) — as well as the bound in (4) — are immediate consequences of Theorem 2.4.

*Proof of Theorem 3.3.* That (A1) holds with the given expressions of  $C_2$  and  $C_4$  is immediate; see Theorem 3.1. Condition (A2) follows by applying Theorem 2.4 to the function

$$F(x_0, \dots, x_{n-K}, x_n, x_{n+K}, \dots, x_{N-1}) = f(x_n) \cdot \nabla h \left( v + \frac{1}{\sqrt{N}} \sum_{k \in [0, N) \setminus [n]_{N,K}} f(x_k) t \right),$$

<sup>2</sup>This is a standard condition, requiring that, given  $u \in \mathbb{R}^d \setminus \{0\}$ , the scalar function  $u \cdot f$  cannot be written in the form  $g - g \circ T$  for any  $L^2(\mu)$  function  $g : X \rightarrow \mathbb{R}$ .



and two index sets  $I_1 \leq I_2$ , where either  $I_1 = (0, \dots, n-K, n)$  and  $I_2 = (n+K, \dots, N-1)$  (case  $n \geq K$ ), or  $I_1 = (0, \dots, n-K)$  and  $I_2 = (n, n+K, \dots, N-1)$  (case  $n < K$ ). The function  $x_n \mapsto F(x_0, \dots, x_{N-1})$  belongs to

$$\mathcal{H}_-(d\|\nabla h\|_\infty c, \vartheta) \cap \mathcal{H}_+(d\|\nabla h\|_\infty c, \vartheta),$$

and for other indices  $r \neq n$ , the function  $x_r \mapsto F(x_0, \dots, x_{N-1})$  belongs to

$$\mathcal{H}_-(N^{-\frac{1}{2}}d^2\|f\|_\infty\|D^2h\|_\infty c, \vartheta) \cap \mathcal{H}_+(N^{-\frac{1}{2}}d^2\|f\|_\infty\|D^2h\|_\infty c, \vartheta).$$

Hence, we see that  $F$  is  $(I_1, I_2)$ -admissible with the same parameters

$$d^2c(\|\nabla h\|_\infty + \|f\|_\infty\|D^2h\|_\infty) \quad \text{and} \quad \vartheta.$$

By Theorem 2.4, (A2) is satisfied with the value of  $C_0$  given.  $\square$

**3.3. Multivariate normal approximation by Pène's method.** In [11], Pène introduced a method of multivariate normal approximation based on the work of Rio [13]; see also [9, 10] for earlier, related, results by the same author. The theorem below is a special case of Pène's theorem applied to a map  $T : X \rightarrow X$  preserving a probability measure  $\mu$ . We write

$$S_N = \sum_{k=0}^{N-1} f^k.$$

Otherwise the notation is the same as in the previous section.

**Theorem 3.4.** *Let  $f : X \rightarrow \mathbb{R}^d$  be a bounded measurable function with  $\mu(f) = 0$ . Suppose that there exist  $r \in \mathbb{Z}_+$ ,  $C \geq 1$ ,  $M \geq \max\{1, \|f\|_\infty\}$  and a sequence of non-negative real numbers  $(\varphi_{p,l})_{p,l}$  such that the following conditions hold:*

**(B1)**  $\varphi_{p,l} \leq 1$  and  $\sum_{p=1}^\infty p \max_{0 \leq l \leq \lfloor p/(r+1) \rfloor} \varphi_{p,l} < \infty$ .

**(B2)** *For any integers  $a, b, c$  satisfying  $1 \leq a + b + c \leq 3$ ; for any integers  $i, j, k, p, q, l$  with  $0 \leq i \leq j \leq k \leq k + p \leq k + p + q \leq k + p + l$ ; for any  $\alpha, \beta, \gamma \in \{1, \dots, d\}$ ; and for any bounded differentiable function  $G : \mathbb{R}^d \times ([-M, M]^d)^3 \rightarrow \mathbb{R}$  with bounded gradient,*

$$|\text{Cov}_\mu[G(S_i, f^i, f^j, f^k), (f_\alpha^{k+p})^a (f_\beta^{k+p+q})^b (f_\gamma^{k+p+l})^c]| \leq C(\|G\|_\infty + \|\nabla G\|_\infty) \varphi_{p,l}.$$

Then the limit

$$\Sigma = \lim_{N \rightarrow \infty} \mu(W_N \otimes W_N) \tag{5}$$

exists. If  $\Sigma = 0$ , then the sequence  $(S_N)_{N \geq 0}$  is bounded in  $L^2(\mu)$ . Otherwise there exists  $B > 0$  such that for any Lipschitz continuous function  $h : \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$|\mu(h(W_N)) - \Phi_\Sigma(h)| \leq \frac{B \text{Lip}(h)}{\sqrt{N}}$$

for all  $N \geq 1$ .

We proceed to the case of billiards:

**Theorem 3.5.** *Assume  $f : X \rightarrow \mathbb{R}^d$  is bounded,  $\int f d\mu = 0$ , and there exist constants  $c \geq 0$  and  $\vartheta \in (0, 1)$  such that  $f_\alpha \in \mathcal{H}_-(c, \vartheta) \cap \mathcal{H}_+(c, \vartheta)$  for all  $\alpha \in \{1, \dots, d\}$ . Then conditions (B1) and (B2) are satisfied with  $\varphi_{p,l} = \theta^p$  and*

$$C = M \frac{6d(c+1) \max\{1, \|f\|_\infty^5\}}{1 - \vartheta}.$$

The result is due to Pène [11], assuming piecewise Hölder continuous observables. The above version covers also dynamically Hölder continuous observables. But again, our intention here is to underline that the conditions of Pène's theorem are immediate consequences of Theorem 2.4.

*Proof of Theorem 3.5.* Obviously  $\varphi_{p,l} = \theta^p$  satisfies (B1). To establish condition (B2), define

$$\begin{aligned} F(x_0, \dots, x_{i-1}, x_i, x_j, x_k, x_{k+p}, x_{k+p+q}, x_{k+p+l}) \\ = G\left(\sum_{m=0}^{i-1} f(x_m), f(x_i), f(x_j), f(x_k)\right) f_\alpha^a(x_{k+p}) f_\beta^b(x_{k+p+q}) f_\gamma^c(x_{k+p+l}). \end{aligned}$$

Then  $\|F\|_\infty \leq \|G\|_\infty \|f\|_\infty^{a+b+c}$ , and  $F$  is  $(I_1, I_2)$ -admissible, where  $I_1 = (0, \dots, i-1, i, j, k)$  and  $I_2 = (k+p, k+p+q, k+p+l)$ : For the indices  $r \in \{k+p, k+p+q, k+p+l\}$ , since  $a, b, c$  are integers with  $1 \leq a+b+c \leq 3$ , a simple computation shows that the function  $x_r \mapsto F(x_0, \dots, x_{k+p+l})$  belongs to

$$\mathcal{H}_-(\|G\|_\infty 3 \max\{1, \|f\|_\infty^5\} c, \vartheta) \cap \mathcal{H}_+(\|G\|_\infty 3 \max\{1, \|f\|_\infty^5\} c, \vartheta).$$

Moreover, for  $r \in \{0, \dots, i-1, i, j, k\}$ , the function  $x_r \mapsto F(x_0, \dots, x_{k+p+l})$  is in

$$\mathcal{H}_-(d\|\nabla G\|_\infty \max\{1, \|f\|_\infty^3\} c, \vartheta) \cap \mathcal{H}_+(d\|\nabla G\|_\infty \max\{1, \|f\|_\infty^3\} c, \vartheta).$$

Consequently,  $F$  is  $(I_1, I_2)$ -admissible with the same parameters

$$3dc \max\{1, \|f\|_\infty^5\} (\|G\|_\infty + \|\nabla G\|_\infty) \quad \text{and} \quad \vartheta.$$

Theorem 2.4 applied to  $F$  and  $(I_1, I_2)$  now yields the estimate

$$\begin{aligned} |\text{Cov}_\mu[G(S_i, f^i, f^j, f^k), (f_\alpha^{k+p})^a (f_\beta^{k+p+q})^b (f_\gamma^{k+p+l})^c]| \\ \leq M \left( \frac{3dc \max\{1, \|f\|_\infty^5\} (\|G\|_\infty + \|\nabla G\|_\infty)}{1 - \vartheta} + \|F\|_\infty \right) \theta^p \\ \leq M \left( \frac{3dc \max\{1, \|f\|_\infty^5\} (\|G\|_\infty + \|\nabla G\|_\infty)}{1 - \vartheta} + \|G\|_\infty \|f\|_\infty^{a+b+c} \right) \theta^p \\ \leq M \frac{6d(c+1) \max\{1, \|f\|_\infty^5\}}{1 - \vartheta} (\|G\|_\infty + \|\nabla G\|_\infty) \theta^p. \end{aligned}$$

Hence, (B2) holds with the value of  $C$  given.  $\square$

**3.4. Vector-valued almost sure invariance principle by Gouëzel's method.** In this section we present an application of Theorem 2.4 to multivariate almost sure limits.

The following theorem is due to Gouëzel [4].

**Theorem 3.6.** *Let  $f : X \rightarrow \mathbb{R}^d$  be a bounded measurable function with  $\mu(f) = 0$ . Given integers  $n > 0$ ,  $m > 0$ ,  $0 \leq b_1 < b_2 < \dots < b_{n+m+1}$ ,  $k \geq 0$ , and vectors  $t_1, \dots, t_{n+m} \in \mathbb{R}^d$ , set*

$$X_{n,m}^{(k)} = \sum_{j=n}^m t_j \cdot \sum_{i=b_j+k}^{b_{j+1}-1+k} f^i$$

for brevity. Now, suppose there exist constants  $t > 0$ ,  $C > 0$ ,  $C' > 0$  and  $\theta \in (0, 1)$  such that

$$\left| \mu\left(e^{iX_{1,n}^{(0)} + iX_{n+1,n+m}^{(k)}}\right) - \mu\left(e^{iX_{1,n}^{(0)}}\right) \mu\left(e^{iX_{n+1,n+m}^{(k)}}\right) \right| \leq C\theta^k \left(1 + \max_{1 \leq j \leq n+m} |b_{j+1} - b_j|\right)^{C'(n+m)} \quad (6)$$

holds for all choices of the numbers  $n, m, b_j, k > 0$ , and all vectors  $t_j$  satisfying  $|t_j| < t$ . Then

(1) Equation (3) yields a well-defined, symmetric, positive-semidefinite,  $d \times d$  matrix  $\Sigma$ .

- (2) The matrix  $\Sigma$  satisfies (5).
- (3)  $W_N$  converges in distribution to  $\mathcal{N}(0, \Sigma)$ .
- (4) Given any  $\lambda > \frac{1}{4}$ , there exists a probability space together with two  $\mathbb{R}^d$ -valued processes  $(Y_n)_{n \geq 0}$  and  $(Z_n)_{n \geq 0}$  such that
  - (a)  $(f^n)_{n \geq 0}$  and  $(Y_n)_{n \geq 0}$  have the same distribution.
  - (b) The random vectors  $Z_n \sim \mathcal{N}(0, \Sigma)$  are independent.
  - (c) Almost surely,  $|\sum_{k=0}^{n-1} Y_k - \sum_{k=0}^{n-1} Z_k| = o(n^\lambda)$ .

Such a theorem has a multitude of interesting consequences, including the central limit theorem (CLT), weak invariance principle, almost sure CLT, law of the iterated logarithm (LIL), Strassen's functional LIL, an upper and lower class refinement of the LIL, and an upper and lower class refinement of Chung's LIL. We refer the reader to [1, 6, 8, 12, 17] for more details concerning the implications.

We proceed to check condition (6) in the case of billiards. This was done by direct means in [16]. To our knowledge, the resulting vector-valued almost sure invariance principle comes with the smallest error and covers the broadest class of observables to date. Here we show condition (6) to be an immediate consequence of Theorem 2.4.

**Theorem 3.7.** Assume  $f : X \rightarrow \mathbb{R}^d$  is bounded,  $\int f d\mu = 0$ , and there exist constants  $c \geq 0$  and  $\vartheta \in (0, 1)$  such that  $f_\alpha \in \mathcal{H}_-(c, \vartheta) \cap \mathcal{H}_+(c, \vartheta)$  for all  $\alpha \in \{1, \dots, d\}$ . Given  $t > 0$ ,

$$\left| \mu \left( e^{iX_{1,n}^{(0)} + iX_{n+1,n+m}^{(k)}} \right) - \mu \left( e^{iX_{1,n}^{(0)}} \right) \mu \left( e^{iX_{n+1,n+m}^{(k)}} \right) \right| \leq \sqrt{2}M \left( \frac{t\sqrt{dc}}{1-\vartheta} + 1 \right) \theta^k$$

holds for all choices of the numbers  $n, m, b_j, k > 0$ , and all vectors  $t_j$  satisfying  $|t_j| < t$ .

*Proof.* Let  $I_1 = (b_j, \dots, b_{j+1}-1 : 1 \leq j \leq n)$  and  $I_2 = (b_j+k, \dots, b_{j+1}-1+k : n+1 \leq j \leq n+m)$ . Define the function

$$\begin{aligned} & F(x_{b_1}, \dots, x_{b_2-1}, \dots, x_{b_{n+m}+k}, \dots, x_{b_{n+m+1}-1+k}) \\ &= \exp \left( i \sum_{j=1}^n t_j \cdot \sum_{i=b_j}^{b_{j+1}-1} f(x_i) + i \sum_{j=n+1}^{n+m} t_j \cdot \sum_{i=b_j+k}^{b_{j+1}-1+k} f(x_i) \right). \end{aligned}$$

Then  $F$  is  $(I_1, I_2)$ -admissible with the same parameters  $t\sqrt{dc}$  and  $\vartheta$ . Indeed, for all indices  $r$ ,

$$x_r \mapsto F(x_{b_1}, \dots, x_{b_{n+m+1}-1+k}) \in \mathcal{H}_-(t\sqrt{dc}, \vartheta) \cap \mathcal{H}_+(t\sqrt{dc}, \vartheta).$$

To see this, recall that  $|e^{ia} - e^{ib}| \leq |a - b|$  for all  $a, b \in \mathbb{R}$ . Thus, if say  $r = b_1$  and  $x \in W^u(y)$ ,

$$\begin{aligned} & |F(x, x_{b_1+1}, \dots, x_{b_{n+m+1}-1+k}) - F(y, x_{b_1+1}, \dots, x_{b_{n+m+1}-1+k})| \\ & \leq |t_1 \cdot f(x) - t_1 \cdot f(y)| \leq |t_1| \left( \sum_{1 \leq \alpha \leq d} |f_\alpha(x) - f_\alpha(y)|^2 \right)^{1/2} \leq t\sqrt{dc} \vartheta^{s+(x,y)}. \end{aligned}$$

The other indices and local stable manifolds are treated similarly. Theorem 2.4 now yields

$$\left| \mu \left( e^{iX_{1,n}^{(0)} + iX_{n+1,n+m}^{(k)}} \right) - \mu \left( e^{iX_{1,n}^{(0)}} \right) \mu \left( e^{iX_{n+1,n+m}^{(k)}} \right) \right| \leq \sqrt{2}M \left( \frac{t\sqrt{dc}}{1-\vartheta} + \|F\|_\infty \right) \theta^k.$$

Since  $\|F\|_\infty = 1$ , the proof is complete.  $\square$

## 4. PROOFS OF THEOREMS 2.3 AND 2.4

We begin by recalling three facts from the theory of billiards, which are necessary for the proofs of the theorems. We refer the reader to the standard textbook [3] for more details.

**Lemma 4.1.** *The space  $(X, \text{Borel}, \mu)$  is a standard probability space, and the family  $\xi = \{\xi_q : q \in \mathcal{Q}\}$  of local unstable manifolds is a measurable partition of it. Here  $\mathcal{Q}$  is an uncountable index set. Thus, the measure  $\mu$  admits a disintegration*

$$\mu = \int_{\mathcal{Q}} \nu_q d\lambda(q),$$

where the  $\{\nu_q : q \in \mathcal{Q}\}$  is a system of conditional probability measures of  $\mu$  on  $\xi$ , with  $\nu_q(\xi_q) = 1$  almost surely, and  $\lambda$  is a factor probability measure on  $\mathcal{Q}$ .

**Lemma 4.2.** *There exist system constants  $a_0 > 0$ ,  $M_0 > 0$  and  $\theta_0 \in (0, 1)$  such that the following holds. Suppose  $G \in \mathcal{H}_-(c, \vartheta)$ . Then,*

$$\left| \int_{\xi_q} G \circ T^n d\nu_q - \int G d\mu \right| \leq M_0(c + \|G\|_\infty) \max(\theta_0, \vartheta)^{\frac{1}{2}(n - a_0 \log |\xi_q|)}$$

for all  $n \geq 0$  and  $q \in \mathcal{Q}$ . Here  $|\xi_q|$  stands for the length of the local unstable manifold  $\xi_q$ .

**Lemma 4.3.** *There exists a system constant  $M_1 > 0$  such that*

$$\int_{\mathcal{Q}} |\xi_q|^{-1} d\lambda(q) \leq M_1.$$

Moreover,

$$\lambda(\{q \in \mathcal{Q} : |\xi_q| \leq \varepsilon\}) \leq M_1 \varepsilon$$

for all  $\varepsilon > 0$ .

Next, let us recall a simple lemma.

**Lemma 4.4.** *Let  $F : X^p \rightarrow \mathbb{R}$  be a function, with  $p \geq 1$  arbitrary. Then the identity*

$$F(x_1, \dots, x_p) - F(y_1, \dots, y_p) = \sum_{r=1}^p [F(x_{<r}, x_r, y_{>r}) - F(x_{<r}, y_r, y_{>r})] \quad (7)$$

holds for all  $(x_1, \dots, x_p), (y_1, \dots, y_p) \in X^p$ . Here we have denoted  $x_{<r} = (x_1, \dots, x_{r-1})$  and  $y_{>r} = (y_{r+1}, \dots, y_p)$ , with the agreement that  $x_{<1} = y_{>p} = \emptyset$ .

*Proof.* The claim is tautological for  $p = 1$ . For  $p > 1$ , the induction step

$$\begin{aligned} & F(x_1, \dots, x_p) - F(y_1, \dots, y_p) \\ &= F(x_{<p}, x_p) - F(x_{<p}, y_p) + F(x_{<p}, y_p) - F(y_{<p}, y_p) \\ &= F(x_{<p}, x_p, y_{>p}) - F(x_{<p}, y_p, y_{>p}) + \sum_{r=1}^{p-1} [F(x_{<r}, x_r, y_{>r}) - F(x_{<r}, y_r, y_{>r})] \end{aligned}$$

proves the lemma. □

The next lemma is a reflection of the fact that  $\mathcal{H}_-$  and  $\mathcal{H}_+$  are dynamically closed, as mentioned earlier; see also [15].

**Lemma 4.5.** *Let  $F : X^{p+q} \rightarrow \mathbb{R}$  be dynamically Hölder continuous in the sense that*

$$x_r \mapsto F(x_1, \dots, x_{p+q}) \in \begin{cases} \mathcal{H}_+(c_r, \vartheta_+), & 1 \leq r \leq p, \\ \mathcal{H}_-(c_r, \vartheta_-), & p+1 \leq r \leq p+q. \end{cases}$$

*Let  $i_1 < \dots < i_p \leq 0$  and  $0 \leq j_1 < \dots < j_q$ . Then*

$$x \mapsto F(T^{i_1}x, \dots, T^{i_p}x, y_1, \dots, y_q) \in \mathcal{H}_+\left(\sum_{r=1}^p c_r \vartheta_+^{-i_r}, \vartheta_+\right)$$

*for all  $(y_1, \dots, y_q) \in X^q$  and*

$$y \mapsto F(x_1, \dots, x_p, T^{j_1}y, \dots, T^{j_q}y) \in \mathcal{H}_-\left(\sum_{r=p+1}^{p+q} c_r \vartheta_-^{j_r}, \vartheta_-\right)$$

*for all  $(x_1, \dots, x_p) \in X^p$ .*

*Proof.* Denote  $G(x) = F(T^{i_1}x, \dots, T^{i_p}x, y_1, \dots, y_q)$ . Let  $x$  and  $\bar{x}$  belong to the same local unstable manifold,  $x \in W^u(\bar{x})$ . Recalling  $i_r \leq 0$ , we have  $T^{i_r}x \in W^u(T^{i_r}\bar{x})$ , and  $s_+(T^{i_r}x, T^{i_r}\bar{x}) = s_+(x, \bar{x}) - i_r$ . Thus, identity (7) yields

$$|G(x) - G(\bar{x})| \leq \sum_{r=1}^p c_r \vartheta_+^{s_+(T^{i_r}x, T^{i_r}\bar{x})} = \left(\sum_{r=1}^p c_r \vartheta_+^{-i_r}\right) \vartheta_+^{s_+(x, \bar{x})},$$

proving the first claim. A corresponding result holds for the inverse map  $T^{-1}$ , which is equivalent to the second claim.  $\square$

We are now ready to prove the first theorem, concerning  $K = 2$ .

*Proof of Theorem 2.3.*

$$I_1 = (i_1, i_2, \dots, i_{p_1}) \quad \text{and} \quad I_2 = (i_{p_1+1}, i_{p_1+2}, \dots, i_{p_2}).$$

Since the billiard process is stationary, the finite-dimensional distributions are the same for the translated index sets

$$I'_1 = I_1 - m = (i'_1, i'_2, \dots, i'_{p_1}) \quad \text{and} \quad I'_2 = I_2 - m = (i'_{p_1+1}, i'_{p_1+2}, \dots, i'_{p_2}),$$

where

$$i'_r = i_r - m, \quad 1 \leq r \leq p_2,$$

and  $m \in \mathbb{N}$  is a number to be determined later. For the moment it suffices to assume that  $i_{p_1} < m \leq i_{p_1+1}$ , meaning  $i'_{p_1} < 0 \leq i'_{p_1+1}$ .

For brevity, define

$$G(x, y) = F(T^{i'_1}x, \dots, T^{i'_{p_1}}x, T^{i'_{p_1+1}-i'_{p_1+1}}y, \dots, T^{i'_{p_2}-i'_{p_1+1}}y).$$

Of course, we then have

$$G(y, T^{i'_{p_1+1}}y) = F(T^{i'_1}y, \dots, T^{i'_{p_1}}y, T^{i'_{p_1+1}}y, \dots, T^{i'_{p_2}}y),$$

and

$$\int F \, d\mathbf{P}_{I'_1 \cup I'_2} = \int_Q \int_{\xi_q} G(y, T^{i'_{p_1+1}}y) \, d\nu_q(y) \, d\lambda(q),$$

Since  $i'_1 < \dots < i'_{p_1} < 0$ , Lemma 4.5 implies that

$$x \mapsto G(x, T^{i'_{p_1+1}}y) \in \mathcal{H}_+\left(\sum_{r=1}^p c_r \vartheta_+^{-i'_r}, \vartheta_+\right),$$

so

$$\sup_{y \in \xi_q} \left| \int_{\xi_q} G(x, T^{i'_{p_1+1}} y) d\nu_q(x) - G(y, T^{i'_{p_1+1}} y) \right| \leq \sum_{r=1}^{p_1} c_r \vartheta_+^{-i'_r}.$$

Inserting this estimate into the identity above, we obtain

$$\left| \int F d\mathbf{P}_{I'_1 \cup I'_2} - \int_{\mathcal{Q}} \int_{\xi_q} \int_{\xi_q} G(x, T^{i'_{p_1+1}} y) d\nu_q(y) d\nu_q(x) d\lambda(q) \right| \leq \sum_{r=1}^{p_1} c_r \vartheta_+^{-i'_r}$$

after an application of Fubini's theorem.

Let us denote

$$\hat{\mathcal{Q}} = \lambda(\{q \in \mathcal{Q} : |\xi_q| \leq e^{-\frac{i'_{p_1+1}}{3a_0}}\}),$$

where  $a_0$  is the constant appearing in Lemma 4.2. Note that, by Lemma 4.3,

$$\mu(\hat{\mathcal{Q}}) \leq M_1 e^{-\frac{i'_{p_1+1}}{3a_0}}.$$

Since  $0 = i'_{p_1+1} - i'_{p_1+1} < \dots < i'_{p_2} - i'_{p_1+1}$ , Lemma 4.5 implies that

$$y \mapsto G(x, y) \in \mathcal{H}_- \left( \sum_{r=p_1+1}^{p_2} c_r \vartheta_-^{i'_r - i'_{p_1+1}}, \vartheta_- \right).$$

Therefore, by Lemma 4.2,

$$\begin{aligned} & \left| \int_{\xi_q} G(x, T^{i'_{p_1+1}} y) d\nu_q(y) - \int G(x, y) d\mu(y) \right| \\ & \leq M_0 \left( \sum_{r=p_1+1}^{p_2} c_r \vartheta_-^{i'_r - i'_{p_1+1}} + \|G\|_\infty \right) \max(\theta_0, \vartheta_-)^{\frac{1}{2}(i'_{p_1+1} - \frac{1}{3}i'_{p_1+1})} \\ & \leq M_0 \left( \sum_{r=p_1+1}^{p_2} c_r \vartheta_-^{i'_r - i'_{p_1+1}} + \|F\|_\infty \right) \max(\theta_0, \vartheta_-)^{\frac{1}{3}i'_{p_1+1}} \equiv E \end{aligned}$$

if  $q \in \mathcal{Q} \setminus \hat{\mathcal{Q}}$ . On the other hand, if  $q \in \hat{\mathcal{Q}}$ , we have the trivial bound

$$\left| \int_{\xi_q} G(x, T^{i'_{p_1+1}} y) d\nu_q(y) - \int G(x, y) d\mu(y) \right| \leq 2\|G\|_\infty \leq 2\|F\|_\infty.$$

This yields

$$\begin{aligned} & \left| \int_{\xi_q} \int_{\xi_q} G(x, T^{i'_{p_1+1}} y) d\nu_q(y) d\nu_q(x) - \int_{\xi_q} \int G(x, y) d\mu(y) d\nu_q(x) \right| \\ & \leq E 1_{\mathcal{Q} \setminus \hat{\mathcal{Q}}}(q) + 2\|F\|_\infty 1_{\hat{\mathcal{Q}}}(q). \end{aligned}$$

Integrating the expression inside the absolute value with respect to  $d\lambda(q)$ , we obtain

$$\begin{aligned} & \left| \int_{\mathcal{Q}} \int_{\xi_q} \int_{\xi_q} G(x, T^{i'_{p_1+1}} y) d\nu_q(y) d\nu_q(x) d\lambda(q) \right. \\ & \quad \left. - \int_{\mathcal{Q}} \int_{\xi_q} \int G(x, y) d\mu(y) d\nu_q(x) d\lambda(q) \right| \\ & \leq E\lambda(\mathcal{Q} \setminus \hat{\mathcal{Q}}) + 2\|F\|_\infty \lambda(\hat{\mathcal{Q}}) \leq E + 2\|F\|_\infty M_1 e^{-\frac{i'_{p_1+1}}{3a_0}}. \end{aligned}$$

Hence,

$$\begin{aligned} & \left| \int F d\mathbf{P}_{I'_1 \cup I'_2} - \int_{\mathcal{Q}} \int_{\xi_q} \int G(x, y) d\mu(y) d\nu_q(x) d\lambda(q) \right| \\ & \leq \sum_{r=1}^{p_1} c_r \vartheta_+^{-i'_r} + E + 2\|F\|_{\infty} M_1 e^{-\frac{i'_{p_1+1}}{3a_0}}. \end{aligned}$$

Recalling  $\int_{\mathcal{Q}} \nu_q d\lambda(q) = \mu$ , Fubini's theorem gives

$$\begin{aligned} & \left| \int F d\mathbf{P}_{I'_1 \cup I'_2} - \iint G(x, y) d\mu(x) d\mu(y) \right| \\ & \leq \sum_{r=1}^{p_1} c_r \vartheta_+^{-i'_r} + M_0 \left( \sum_{r=p_1+1}^{p_2} c_r \vartheta_-^{i'_r - i'_{p_1+1}} + \|F\|_{\infty} \right) \max(\theta_0, \vartheta_-)^{\frac{i'_{p_1+1}}{3}} \\ & \quad + 2\|F\|_{\infty} M_1 e^{-\frac{i'_{p_1+1}}{3a_0}}, \end{aligned}$$

where we note that

$$\iint G(x, y) d\mu(x) d\mu(y) = \int F d(\mathbf{P}_{I'_1} \otimes \mathbf{P}_{I'_2})$$

with  $I'_2 = I'_2 - i'_{p_1+1} = (i'_{p_1+1} - i'_{p_1+1}, \dots, i'_{p_2} - i'_{p_1+1})$ . Stationarity guarantees that

$$\mathbf{P}_{I'_1 \cup I'_2} = \mathbf{P}_{I_1 \cup I_2}, \quad \mathbf{P}_{I'_1} = \mathbf{P}_{I_1} \quad \text{and} \quad \mathbf{P}_{I'_2} = \mathbf{P}_{I_2}.$$

Inserting  $i'_r = i_r - m$  into the upper bound above, we see that

$$\begin{aligned} & \left| \int F d\mathbf{P}_{I_1 \cup I_2} - \int F d(\mathbf{P}_{I_1} \otimes \mathbf{P}_{I_2}) \right| \\ & \leq \sum_{r=1}^{p_1} c_r \vartheta_+^{i_{p_1} - i_r} \vartheta_+^{m - i_{p_1}} + M_0 \left( \sum_{r=p_1+1}^{p_2} c_r \vartheta_-^{i_r - i_{p_1+1}} + \|F\|_{\infty} \right) \max(\theta_0, \vartheta_-)^{\frac{i_{p_1+1} - m}{3}} \\ & \quad + 2\|F\|_{\infty} M_1 e^{-\frac{i_{p_1+1} - m}{3a_0}}, \end{aligned}$$

whenever  $i_{p_1} < m \leq i_{p_1+1}$ . Finally, let  $m$  be the smallest integer  $\geq \frac{1}{4}(3i_{p_1} + i_{p_1+1})$ . Then  $\frac{1}{3}(i_{p_1+1} - m) \geq \frac{1}{4}(i_{p_1+1} - i_{p_1}) - \frac{1}{3}$  and  $m - i_{p_1} \geq \frac{1}{4}(i_{p_1+1} - i_{p_1})$ . This yields the final estimate

$$\begin{aligned} & \left| \int F d\mathbf{P}_{I_1 \cup I_2} - \int F d(\mathbf{P}_{I_1} \otimes \mathbf{P}_{I_2}) \right| \\ & \leq \left( \sum_{r=1}^{p_1} c_r \vartheta_+^{i_{p_1} - i_r} \right) \vartheta_+^{\frac{1}{4}(i_{p_1+1} - i_{p_1})} \\ & \quad + M_0 \left( \sum_{r=p_1+1}^{p_2} c_r \vartheta_-^{i_r - i_{p_1+1}} + \|F\|_{\infty} \right) \max(\theta_0, \vartheta_-)^{\frac{1}{4}(i_{p_1+1} - i_{p_1}) - \frac{1}{3}} \\ & \quad + 2\|F\|_{\infty} M_1 e^{-\frac{1}{4a_0}(i_{p_1+1} - i_{p_1}) + \frac{1}{3a_0}}. \end{aligned}$$

Defining the system constant  $\theta_1 = e^{-\frac{1}{a_0}}$ , we arrive at the claimed bound.  $\square$

We proceed to the proof of the second theorem, concerning  $K \geq 2$ .

*Proof of Theorem 2.4.* The proof is based on induction with respect to  $K$ .



Case  $K = 2$ : The assumption is that  $F$  is  $(I_1, I_2)$ -admissible with the same parameters  $c, \vartheta$ . Therefore, Theorem 2.3 yields

$$\begin{aligned}
& \left| \int F d\mathbf{P}_{I_1 \cup I_2} - \int F d(\mathbf{P}_{I_1} \otimes \mathbf{P}_{I_2}) \right| \\
& \leq \left( c \sum_{r=1}^{p_1} \vartheta^{i_{p_1} - i_r} \right) \vartheta^{\frac{1}{4}\ell_1} + M_0 \left( c \sum_{r=p_1+1}^{p_2} \vartheta^{i_r - i_{p_1+1}} + \|F\|_\infty \right) \max(\theta_0, \vartheta)^{\frac{1}{4}\ell_1 - \frac{1}{3}} \\
& \quad + 2M_1 \|F\|_\infty \theta_1^{\frac{1}{4}\ell_1 - \frac{1}{3}} \\
& \leq \frac{c}{1 - \vartheta} \vartheta^{\frac{1}{4}\ell_1} + M_0 \left( \frac{c}{1 - \vartheta} + \|F\|_\infty \right) \theta_0^{-\frac{1}{3}} \max(\theta_0, \vartheta)^{\frac{1}{4}\ell_1} + 2M_1 \|F\|_\infty \theta_1^{\frac{1}{4}\ell_1 - \frac{1}{3}} \\
& \leq \left( \frac{c}{1 - \vartheta} + M_0 \left( \frac{c}{1 - \vartheta} + \|F\|_\infty \right) \theta_0^{-\frac{1}{3}} + 2M_1 \|F\|_\infty \theta_1^{-\frac{1}{3}} \right) \max(\vartheta, \theta_0, \theta_1)^{\frac{1}{4}\ell_1}.
\end{aligned}$$

Defining the system constants  $M = \max(1 + M_0 \theta_0^{-\frac{1}{3}}, M_0 \theta_0^{-\frac{1}{3}} + 2M_1 \theta_1^{-\frac{1}{3}})$  and  $\theta = \max(\vartheta, \theta_0, \theta_1)^{\frac{1}{4}}$ , we obtain

$$\left| \int F d\mathbf{P}_I - \int F d(\mathbf{P}_{I_1} \otimes \mathbf{P}_{I_2}) \right| \leq M \left( \frac{c}{1 - \vartheta} + \|F\|_\infty \right) \theta^{\ell_1}$$

as claimed.

Case  $K > 2$ : We are now assuming that  $F$  is  $(I_1, \dots, I_K)$ -admissible with the same parameters  $c, \vartheta$ . In particular,  $F$  is then  $(I_1 \cup \dots \cup I_{K-1}, I_K)$ -admissible. Hence, the preceding case implies

$$\left| \int F d\mathbf{P}_{I_1 \cup \dots \cup I_K} - \int F d(\mathbf{P}_{I_1 \cup \dots \cup I_{K-1}} \otimes \mathbf{P}_{I_K}) \right| \leq M \left( \frac{c}{1 - \vartheta} + \|F\|_\infty \right) \theta^{\ell_{K-1}}.$$

Suppose that

$$\left| \int G d\mathbf{P}_{I_1 \cup \dots \cup I_{K-1}} - \int G d(\mathbf{P}_{I_1} \otimes \dots \otimes \mathbf{P}_{I_{K-1}}) \right| \leq M \left( \frac{c}{1 - \vartheta} + \|G\|_\infty \right) \sum_{k=1}^{K-2} \theta^{\ell_k}$$

for all  $(I_1, \dots, I_{K-1})$ -admissible functions  $G$  with the same parameters  $c, \vartheta$ . It now suffices to just observe that  $F$  is  $(I_1, \dots, I_{K-1})$ -admissible in its first  $p_{K-1}$  arguments. More precisely, given  $\mathbf{y} \in X^{p_K - p_{K-1}}$ , the function

$$G_{\mathbf{y}} : X^{p_{K-1}} \rightarrow \mathbb{R} : G_{\mathbf{y}}(\mathbf{x}) = F(\mathbf{x}, \mathbf{y})$$

is  $(I_1, \dots, I_{K-1})$ -admissible, bounded by  $\|F\|_\infty$ . Hence,

$$\begin{aligned}
& \left| \int F d(\mathbf{P}_{I_1 \cup \dots \cup I_{K-1}} \otimes \mathbf{P}_{I_K}) - \int F d(\mathbf{P}_{I_1} \otimes \dots \otimes \mathbf{P}_{I_K}) \right| \\
& = \left| \int \left( \int G_{\mathbf{y}} d\mathbf{P}_{I_1 \cup \dots \cup I_{K-1}} - \int G_{\mathbf{y}} d(\mathbf{P}_{I_1} \otimes \dots \otimes \mathbf{P}_{I_{K-1}}) \right) d\mathbf{P}_{I_K}(\mathbf{y}) \right| \\
& \leq M \left( \frac{c}{1 - \vartheta} + \|F\|_\infty \right) \sum_{k=1}^{K-2} \theta^{\ell_k}.
\end{aligned}$$

This finishes the proof. □

## REFERENCES

- [1] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication. doi:10.1002/9780470316962.
- [2] Nikolai Chernov. Advanced statistical properties of dispersing billiards. *J. Stat. Phys.*, 122(6):1061–1094, 2006. doi:10.1007/s10955-006-9036-8.
- [3] Nikolai Chernov and Roberto Markarian. *Chaotic billiards*, volume 127 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.
- [4] Sébastien Gouëzel. Almost sure invariance principle for dynamical systems by spectral methods. *Ann. Probab.*, 38(4):1639–1671, 2010. doi:10.1214/10-AOP525.
- [5] Olli Hella, Juho Leppänen, and Mikko Stenlund. Stein’s method for dynamical systems. 2016. Preprint. Available from: <https://arxiv.org/abs/1701.02966>.
- [6] Michael T. Lacey and Walter Philipp. A note on the almost sure central limit theorem. *Statist. Probab. Lett.*, 9(3):201–205, 1990. doi:10.1016/0167-7152(90)90056-D.
- [7] Juho Leppänen. Functional correlation decay and multivariate normal approximation for non-uniformly expanding maps. 2017. Preprint. Available from: <https://arxiv.org/abs/1702.00699>.
- [8] Ian Melbourne and Matthew Nicol. A vector-valued almost sure invariance principle for hyperbolic dynamical systems. *Ann. Probab.*, 37(2):478–505, 2009. doi:10.1214/08-AOP410.
- [9] Françoise Pène. Rates of convergence in the CLT for two-dimensional dispersive billiards. *Comm. Math. Phys.*, 225(1):91–119, 2002. doi:10.1007/s002201000573.
- [10] Françoise Pène. Multiple decorrelation and rate of convergence in multidimensional limit theorems for the Prokhorov metric. *Ann. Probab.*, 32(3B):2477–2525, 2004. doi:10.1214/009117904000000036.
- [11] Françoise Pène. Rate of convergence in the multidimensional central limit theorem for stationary processes. Application to the Knudsen gas and to the Sinai billiard. *Ann. Appl. Probab.*, 15(4):2331–2392, 2005. doi:10.1214/105051605000000476.
- [12] Walter Philipp and William Stout. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* 2, (issue 2, 161):iv+140, 1975.
- [13] Emmanuel Rio. Sur le théorème de Berry-Esseen pour les suites faiblement dépendantes. *Probab. Theory Related Fields*, 104(2):255–282, 1996. doi:10.1007/BF01247840.
- [14] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971)*, Vol. II: Probability theory, pages 583–602. Univ. California Press, Berkeley, Calif., 1972. Available from: <http://projecteuclid.org/euclid.bsmsp/1200514239>.
- [15] Mikko Stenlund. A strong pair correlation bound implies the CLT for Sinai billiards. *J. Stat. Phys.*, 140(1):154–169, 2010. doi:10.1007/s10955-010-9987-7.
- [16] Mikko Stenlund. A vector-valued almost sure invariance principle for Sinai billiards with random scatterers. *Commun. Math. Phys.*, 325(3):879–916, 2014. doi:10.1007/s00220-013-1870-3.
- [17] Volker Strassen. An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 3:211–226 (1964), 1964.

(Juho Leppänen) DEPARTMENT OF MATHEMATICS AND STATISTICS, P.O. BOX 68, FIN-00014 UNIVERSITY OF HELSINKI, FINLAND.

E-mail address: juho.leppanen@helsinki.fi

(Mikko Stenlund) DEPARTMENT OF MATHEMATICS AND STATISTICS, P.O. BOX 35, FIN-40014 UNIVERSITY OF JYVÄSKYLÄ, FINLAND.

E-mail address: mikko.s.stenlund@jyu.fi